Exponential transform of quadratic functional and multiplicative ergodicity of a Gauss-Markov process

Marina Kleptsyna, Alain Le Breton, Bernard Ycart

Exponential transform of quadratic functional and multiplicative ergodicity of a Gauss-Markov process[☆]

Marina Kleptsyna[a], Alain Le Breton[b], Bernard Ycart[b,c,*]

[a] *Laboratoire Manceau de Mathématiques CNRS EA 3263, Université du Maine, Avenue Olivier Messiaen, 72085 Le Mans cedex 9, France*
[b] *Laboratoire Jean Kuntzmann CNRS UMR 5224, Université Grenoble-Alpes, 51 rue des Mathématiques 38041 Grenoble cedex 9, France*
[c] *Laboratoire d'Excellence TOUCAN (Toulouse Cancer), Toulouse, France*

Abstract

The Laplace transform of partial sums of the square of a non-centered Gauss-Markov process, conditioning on its starting point, is explicitly computed. The parameters of multiplicative ergodicity are deduced.

Keywords: Multiplicative ergodicity, Laplace transform, Gauss-Markov processes

2010 MSC: 60J05, 44A10

[☆] Dedicated to the memory of Michel Viot

[*] corresponding author

Email addresses: Marina.Kleptsyna@univ-lemans.fr (Marina Kleptsyna), Alain.Le-Breton@imag.fr (Alain Le Breton), Bernard.Ycart@imag.fr (Bernard Ycart)

1. Introduction

Asymptotics of exponential transforms for partial sums of functionals of a Markov chain are usually described by multiplicative ergodicity properties, which have been thoroughly studied by Meyn and his co-workers [1, 2, 3, 4]; see also [5, p. 519] for a short introduction. Different presentations of the notion, at increasingly sharp levels, can be given. Here is the definition that will be adopted in this paper, the notations being compatible with those of [3].

Definition 1. *Let $\{X_t\,,\ t \in \mathbb{N}\}$ be a discrete time stochastic process, taking values in a Polish state space $\mathbf{X}$. Let F be a measurable functional from the state space $\mathbf{X}$ into $\mathbb{R}$. For $t \in \mathbb{N}$, denote by S_t the partial sum:*

$$S_t = \sum_{s=0}^{t} F(X_s)\ .$$

The process $\{X_t\,,\ t \in \mathbb{N}\}$ is said to be multiplicatively ergodic for F if there exist a non-empty open subset D of $\mathbb{C}$, a function $\alpha \mapsto \Lambda(\alpha)$ from D into $\mathbb{R}$, and a function $(\alpha, x) \mapsto \check{f}(\alpha, x)$ from $D \times \mathbf{X}$ into $\mathbb{R}^+$, such that for all $\alpha \in D$ and all $x \in \mathbf{X}$:

$$\lim_{t\to\infty} \mathbb{E}_x \left[\exp\left(\alpha S_t - t\Lambda(\alpha)\right)\right] = \check{f}(\alpha, x)\ , \tag{1}$$

where $\mathbb{E}_x$ denotes the conditional expectation given $X_0 = x$.

Balaji and Meyn [1] introduced the notion for a Markov chain on a countable state space. It was later extended to Markov chains on general state spaces by Kontoyannis and Meyn [2, 3]. Applications to large deviations of Markov chains and Monte-Carlo simulation were developed by Meyn [4]. Since then, the notion does not seem to have attracted much further attention. One of the reasons may be that, except in the trivial case where the X_t's are independent, explicit calculations of the eigenvalue $\Lambda(\alpha)$ and the eigenfunction $\check{f}(\alpha, x)$ remain out of reach. Recently, in the study of an exponential cell growth model [6], the necessity of an explicit determination of the constants of multiplicative ergodicity $\Lambda(\alpha)$ and $\check{f}(\alpha, x)$, was highlighted.

The main result of this note (Proposition 2) is an explicit expression for the exponential transform $\mathbb{E}_x[\exp(\alpha S_t)]$, in the particular case where $\{X_t\,,\ t \in \mathbb{N}\}$ is a real-valued, non-centered, stationary autoregressive process, and $F(x) = x^2$. The multiplicative ergodicity coefficients $\Lambda(\alpha)$ and $\check{f}(\alpha, x)$ are deduced (Corollary 3). Moreover, the convergence in (1) is shown to be exponentially fast. The calculation technique used for Proposition 2 was developed in [7], and Proposition 2 generalizes some of the examples given in that reference. Section 2 contains the notations and statement of the main results, proofs are given in section 3.

2. Notations and statement

The process considered here is a stationary autoregressive process, classically defined as follows. Let θ be a real such that $0 < |\theta| < 1$. Let $(e_t)_{t\geqslant 1}$ be a sequence of i.i.d.r.v.'s each following the standard Gaussian distribution. Let Y_0, independent from the sequence $(e_t)_{t\geqslant 1}$, following the normal $\mathcal{N}(0,(1-\theta^2)^{-1})$ distribution. For all $t \geqslant 1$ let:

$$Y_t = \theta Y_{t-1} + e_t \ .$$

Then $\{Y_t\,,\ t\in\mathbb{N}\}$ is a stationary centered auto-regressive process. Denoting by m a fixed real, we consider the non-centered process $\{X_t\,,\ t\in\mathbb{N}\}$, with $X_t = Y_t + m$. Let F be the function $x \mapsto x^2$. Define:

$$L_t(\alpha, x) = \mathbb{E}_x[\exp(\alpha S_t)] = \mathbb{E}_x\left[\exp \alpha \left(\sum_{s=0}^{t} F(X_s)\right)\right] \ .$$

The particular case $m = 0$ was treated in [7], example 4.1. The technique used here is similar. Some notations are needed.

The explicit expression of $L_t(\alpha, x)$ uses the two roots of the following equation in λ:

$$\lambda^2 - (-2\alpha + \theta^2 + 1)\lambda + \theta^2 = 0 \ . \tag{2}$$

Assume first that α is real and negative. The two roots are:

$$\lambda_\pm(\alpha) = \frac{-2\alpha + 1 + \theta^2 \pm \sqrt{(-2\alpha + (\theta+1)^2)(-2\alpha + (\theta-1)^2)}}{2} \ . \tag{3}$$

The following inequalities hold:

$$0 < \frac{\lambda_-(\alpha)}{|\theta|} < 1 < \frac{\lambda_+(\alpha)}{|\theta|} \ . \tag{4}$$

The two functions $\lambda_\pm(\alpha)$ admit a maximal analytic extension over an open domain D of $\mathbb{C}$, containing $(-\infty\,;\,0)$, over which the same inequalities hold for their modules.

$$D = \left\{ \alpha \in \mathbb{C}\,,\ 0 < \frac{|\lambda_-(\alpha)|}{|\theta|} < 1 < \frac{|\lambda_+(\alpha)|}{|\theta|} \right\} \ . \tag{5}$$

In what follows, the variable is omitted in $\lambda_\pm = \lambda_\pm(\alpha)$. Let:

$$\beta_+ = \frac{1-\lambda_-}{\lambda_+ - \lambda_-}\ ; \quad \beta_- = \frac{\lambda_+ - 1}{\lambda_+ - \lambda_-}\ , \tag{6}$$

$$\pi_t = \beta_+ \lambda_+^{t+1} + \beta_- \lambda_-^{t+1}\ ;\ \psi_t = \beta_+ \left(\frac{\lambda_+}{\theta}\right)^t + \beta_- \left(\frac{\lambda_-}{\theta}\right)^t \ . \tag{7}$$

Proposition 2. *Let:*

$$\begin{aligned} \nu &= \frac{m(1-\theta)}{-2\alpha + (1-\theta)^2}\ ;\ A = m(1-\theta)\nu\ ; \\ B &= \frac{\theta}{-2\alpha}\left(x - (1-\theta)\nu\right)^2 - \theta\nu^2\ ;\ C = 2\nu(x - (1-\theta)\nu)\ . \end{aligned}$$

Then

$$L_t(\alpha, x) = (\pi_t)^{-1/2}\ \exp\left(\alpha \Sigma_t\right)\ , \tag{8}$$

with

$$\Sigma_t = At + x^2 + B\left(\theta - \frac{\psi_t}{\psi_{t+1}}\right) + C\left(\theta - \frac{1}{\psi_{t+1}}\right)\ . \tag{9}$$

Once an explicit expression of $L_t(\alpha, x)$ has been obtained, deriving its asymptotics as t tends to infinity is easy, using (4). From the same inequalities, it follows that the convergence in (1) holds at exponential speed $O((\theta/\lambda_+)^t)$.

Corollary 3. *For all $m \in \mathbb{R}$, the process $\{X_t\,,\ t \in \mathbb{N}\}$ is multiplicatively ergodic for F in the sense of Definition 1. The domain D is defined by (5). For $\alpha \in D$ and $x \in \mathbb{R}$, the limit (1) holds with*

$$\Lambda(\alpha) = \frac{\alpha m^2(1-\theta)^2}{(-2\alpha + (1-\theta)^2)} - \frac{1}{2}\log(\lambda_+)\ , \tag{10}$$

and

$$\check{f}(\alpha, x) = (\beta_+\lambda_+)^{-1/2} \exp\left(\alpha\left(x^2 + B\left(\theta - \frac{\theta}{\lambda_+}\right) + C\theta\right)\right)\ . \tag{11}$$

In (10), log denotes an analytic extension to D of the ordinary logarithm on $\mathbb{R}^+$ (recall that λ_+ is positive for $\alpha \in (-\infty\,;0\,) \subset D$).

3. Proof

The technique of proof is an application of Theorem 1 in [7]. As in that reference, we shall rewrite the exponential transform $\mathbb{E}_x[\exp(\alpha S_t)]$ as a Laplace transform, and use $\mu/2$ as its variable: in what follows, α is replaced by $\mu = -2\alpha$, and $L_t(\alpha, x)$ by $L_t(\mu, x)$. Since by definition $X_t = Y_t + m$, the initial condition $X_0 = x$ amounts to $Y_0 = x - m$. Hence the conditional distribution of $\{X_t\,,\, t \in \mathbb{N}\}$ given $X_0 = x$ coincides with the distribution of $\{Y_t^x + m\,,\ t \in \mathbb{N}\}$, where $\{Y_t^x\,,\ t \in \mathbb{N}\}$ is defined by $Y_0^x = x - m$, and for all

$t \geqslant 1$,

$$Y_t^x = \theta Y_{t-1}^x + e_t \ .$$

So our aim is to compute

$$L_t(\mu, x) = \mathbb{E}\left[\exp\left(-\frac{\mu}{2}\sum_{s=0}^{t}(Y_s^x + m)^2\right)\right] \ .$$

The process $\{Y_t^x + m\ ,\ t \geqslant 0\}$ has the same covariance function as $\{Y_t^x\ ,\ t \geqslant 0\}$. Denote by m_t^x its mean function. It is such that $m_0^x = x$ and for $s \geqslant 1$,

$$m_s^x = \theta m_{s-1}^x + m(1-\theta) \ . \tag{12}$$

Theorem 1 of [7], and the calculations in Example 4.1 therein, yield (8), where π_t is defined by (7), and

$$\Sigma_t = \sum_{s=0}^{t} \frac{\psi_s}{\theta\psi_{s+1}}(z_s)^2 \ , \tag{13}$$

where $z_0 = m_0^x = x$ and

$$z_s = m_s^x - \sum_{r=0}^{s-1} \theta^{s-r}\left(1 - \frac{\psi_r}{\theta\psi_{r+1}}\right) z_r \ . \tag{14}$$

Using the expression of m_s^x (12):

$$\begin{aligned} z_s &= (1-\theta)m + \theta m_{s-1}^x - \theta\left(1 - \frac{\psi_{s-1}}{\theta\psi_s}\right) z_{s-1} - \theta\sum_{r=0}^{s-2} \theta^{s-1-r}\left(1 - \frac{\psi_r}{\theta\psi_{r+1}}\right) z_r \\ &= (1-\theta)m + \theta z_{s-1} - \theta\left(1 - \frac{\psi_{s-1}}{\theta\psi_s}\right) z_{s-1} = \frac{\psi_{s-1}}{\psi_s} z_{s-1} + (1-\theta)m \ . \end{aligned}$$

The resolvent of that equation is $(\psi_s)^{-1}$, from which the following expression is obtained.

$$z_s = \psi_s^{-1} \left(x + m(1-\theta) \sum_{l=1}^{s} \psi_l \right) \ .$$

Plugging into (13) yields:

$$\Sigma_t = \sum_{s=0}^{t} \frac{1}{\theta \psi_s \psi_{s+1}} \left(x + m(1-\theta) \sum_{l=1}^{s} \psi_l \right)^2 \ . \qquad (15)$$

Setting $z = \lambda_+/\theta$, $z^{-1} = \lambda_-/\theta$, one has $\psi_s = \beta_+ z^s + \beta_- z^{-s}$, and

$$\sum_{l=1}^{s} \psi_l = \beta_+ \frac{z}{1-z}(1-z^s) + \beta_- \frac{z^{-1}}{1-z^{-1}}(1-z^{-s}) \ .$$

Define

$$\Delta_s = x + m(1-\theta) \sum_{l=1}^{s} \psi_l = a_+ z^s + a_- z^{-s} + a \ .$$

with

$$a_+ = m(\theta-1)\beta_+ \frac{z}{1-z} \ ; \quad a_- = m(\theta-1)\beta_- \frac{z^{-1}}{1-z^{-1}} \ ,$$

and $a = x - (a_+ + a_-)$. For $s \geqslant 1$,

$$\psi_{s+1}\psi_{s-1} - (\psi_s)^2 = \beta_+\beta_-(z - z^{-1})^2 \ ,$$

hence

$$\frac{\psi_{s-1}}{\psi_s} - \frac{\psi_s}{\psi_{s+1}} = \frac{\beta_+\beta_-(z-z^{-1})^2}{\psi_s \psi_{s+1}} \ .$$

Now let us choose the three constants A, B, C such that:

$$\frac{(\Delta_s)^2}{\theta \psi_s \psi_{s+1}} = A + B \left(\frac{\psi_{s-1}}{\psi_s} - \frac{\psi_s}{\psi_{s+1}} \right) + C \left(\frac{1}{\psi_s} - \frac{1}{\psi_{s+1}} \right) \ . \qquad (16)$$

For that we need:

$$(\Delta_s)^2 = A\theta\psi_s\psi_{s+1} + B\theta\beta_+\beta_-(z - z^{-1})^2 + C\theta(\psi_{s+1} - \psi_s) \ ,$$

then

$$\begin{aligned}(a_+z^s + a_-z^{-s} + a)^2 &= A\theta(\beta_+z^s + \beta_-z^{-s})(\beta_+z^{s+1} + \beta_-z^{-s-1}) \\ &\quad +B\theta\beta_+\beta_-(z - z^{-1})^2 + C\theta(\beta_+(z-1)z^s + \beta_-(z^{-1} - 1)z^{-s}) \ .\end{aligned}$$

The expressions of A, B, C can be guessed by identifying powers of z in the expression above. The following constants satisfy the requirements.

$$\begin{aligned} A &= \frac{(a_+)^2}{\theta(\beta_+)^2 z} \ , \\ B &= \frac{a^2 + 2a_+a_- - A\theta\beta_+\beta_-(z + z^{-1})}{\theta\beta_+\beta_-(z - z^{-1})^2} \ , \\ C &= \frac{2aa_+}{2\beta_+(z-1)} \ . \end{aligned}$$

For these constants, plugging (16) into (15) and summing yields:

$$\Sigma_t = \frac{x^2}{\theta\psi_0\psi_1} + At + B\left(\frac{\psi_0}{\psi_1} - \frac{\psi_t}{\psi_{t+1}}\right) + C\left(\frac{1}{\psi_1} - \frac{1}{\psi_{t+1}}\right) \ . \tag{17}$$

Substituting $\psi_0 = 1$ and $\psi_1 = \theta^{-1}$ in (17) gives (9).

To finish the proof of Proposition 2, more explicit expressions of A, B, and C must be obtained. As a preliminary observation, recall that $z = \lambda_+/\theta$ and

$z^{-1} = \lambda_-/\theta$ are the two roots of $z^2 - \theta^{-1}(\mu + \theta^2 + 1)z + 1 = 0$. From there the following symmetric functions of the two roots are obtained.

$$\begin{aligned} z + z^{-1} &= \theta^{-1}(\mu + \theta^2 + 1) , \\ (z - z^{-1})^2 &= \theta^{-2}(\mu + (1-\theta)^2)(\mu + (1+\theta)^2) , \\ (z-1)(1-z^{-1}) &= z + z^{-1} - 2 = \theta^{-1}(\mu + (1-\theta)^2) , \\ (z+1)(1+z^{-1}) &= z + z^{-1} + 2 = \theta^{-1}(\mu + (1+\theta)^2) . \end{aligned}$$

The constants β_+ and β_- can be written as functions of z:

$$\beta_+ = \frac{1 - \theta z^{-1}}{\theta(z - z^{-1})} \quad \text{and} \quad \beta_- = \frac{\theta z - 1}{\theta(z - z^{-1})} .$$

From there:

$$\begin{aligned} \beta_+\beta_- &= \frac{(1-\theta z^{-1})(\theta z - 1)}{\theta^2(z - z^{-1})^2} = \frac{\theta(z + z^{-1}) - 1 - \theta^2}{\theta^2(z - z^{-1})^2} \\ &= \frac{\mu + \theta^2 + 1 - 1 - \theta^2}{\theta^2(z - z^{-1})^2} = \frac{\mu}{(\mu + (1-\theta)^2)(\mu + (1-\theta)^2)} . \end{aligned}$$

Now:

$$\begin{aligned} A &= \frac{a_+^2}{\theta\beta_+^2 z} = \frac{m^2(1-\theta)^2 z(1-z^{-1})^2}{\theta((z-1)(1-z^{-1}))^2} \\ &= \frac{m^2(1-\theta)^2(z + z^{-1} - 2)}{\theta((z-1)(1-z^{-1}))^2} = \frac{m^2(1-\theta)^2}{\mu + (1-\theta)^2} = m(1-\theta)\nu . \end{aligned}$$

Here is the calculation of B:

$$B = \frac{a^2 + 2a_+a_- - A\theta\beta_+\beta_-(z + z^{-1})}{\theta\beta_+\beta_-(z - z^{-1})^2} = \frac{N}{D} .$$

From the expression of $\beta_+\beta_-$ above, $D = \mu/\theta$. The numerator N contains three terms. The first term is:

$$a^2 = (x - (a_+ + a_-))^2 \; ,$$

where,

$$\begin{aligned} a_+ + a_- &= -m(1-\theta)\left(\beta_+ \frac{z}{1-z} + \beta_- \frac{z^{-1}}{1-z^{-1}}\right) \; , \\ &= -m(1-\theta)\frac{\beta_+ z(1-z^{-1}) + \beta_- z^{-1}(1-z)}{(1-z)(1-z^{-1})} \; , \\ &= -m(1-\theta)\frac{\psi_1 - \psi_0}{(1-z)(1-z^{-1})} \; , \\ &= \frac{m(1-\theta)^2}{\mu + (1-\theta)^2} = (1-\theta)\nu \; . \end{aligned}$$

Thus:

$$a^2 = (x - (1-\theta)\nu)^2 \; .$$

The second term in the numerator N is:

$$\begin{aligned} 2a_+a_- &= \frac{2m^2(1-\theta)^2\beta_+\beta_-}{(1-z)(1-z^{-1})} \\ &= \frac{2m^2(1-\theta)^2(-\theta\mu)}{(\mu + (1+\theta)^2)(\mu + (1-\theta)^2)^2} \\ &= -\frac{2\theta\mu}{\mu + (1+\theta)^2}\,\nu^2 \; . \end{aligned}$$

The last term in the numerator N is:

$$
\begin{aligned}
-\theta A \beta_+ \beta_- (z + z^{-1}) &= -\theta A \frac{\mu(\mu + \theta^2 + 1)}{(\mu + (1+\theta)^2)(\mu + (1-\theta)^2)} \\
&= -\frac{m^2(1-\theta)^2 \mu(\mu + \theta^2 + 1)}{(\mu + (1+\theta)^2)(\mu + (1-\theta)^2)^2} \\
&= -\frac{\mu(\mu + \theta^2 + 1)}{\mu + (1+\theta)^2}\, \nu^2 \ .
\end{aligned}
$$

Grouping the three terms and multiplying by $D^{-1} = \theta/\mu$:

$$
\begin{aligned}
B &= \frac{\theta}{\mu}(x - (1-\theta)\nu)^2 - \frac{2\theta^2 + \theta(\mu + \theta^2 + 1)}{\mu + (1+\theta)^2}\, \nu^2 \ , \\
&= \frac{\theta}{\mu}(x - (1-\theta)\nu)^2 - \theta\nu^2 \ .
\end{aligned}
$$

Finally, here is the calculation of C.

$$
\begin{aligned}
C &= \frac{2aa_+}{\theta\beta_+(z-1)} \\
&= 2(x - (1-\theta)\nu) \times \frac{m(1-\theta)z}{\theta(z-1)^2} \ ,
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{z}{\theta(z-1)^2} &= \frac{z(1-z^{-1})^2}{\theta((z-1)(1-z^{-1}))^2} \\
&= \frac{z + z^{-1} - 2}{\theta((z-1)(1-z^{-1}))^2} \\
&= \frac{1}{\mu + (1-\theta)^2} \ .
\end{aligned}
$$

Therefore:

$$C = 2\nu(x - (1-\theta)\nu) \ .$$

This ends the proof of Proposition 2.

Concluding remarks

(a) The (non-conditional) exponential transform $\mathbb{E}[\exp(\alpha S_t)]$ can be obtained by integrating the right-hand side of (8) against the centered Gaussian distribution with variance $(1-\theta^2)^{-1}$. A simpler calculation can be carried through, applying again Theorem 1 of [7] as was done above. Similar explicit expressions are obtained, that will not be reproduced here.

(b) Here, the parameters of multiplicative ergodicity have been deduced from the explicit expression of the Laplace transform, for a specific class of Gaussian processes. On the basis of the results of [7], there is some hope that they could be directly calculated for more general processes. This should be investigated in a forthcoming paper.

[1] S. Balaji, S. P. Meyn, Multiplicative ergodicity and large deviations for an irreducible Markov chain, Stoch. Proc. Appl. 90 (1) (2000) 123–144.

[2] I. Kontoyiannis, S. P. Meyn, Spectral theory and limit theorems for geometrically ergodic Markov processes, Ann. Appl. Probab. 13 (1) (2003) 304–362.

[3] I. Kontoyiannis, S. P. Meyn, Large deviation asymptotics and the spectral theory of multiplicatively regular Markov processes, Elect. J. Probab. 10 (3) (2005) 61–123.

[4] S. P. Meyn, Large deviation asymptotics and control variates for simulating large functions, Ann. Appl. Probab. 16 (1) (2006) 310–339.

[5] S. P. Meyn, R. L. Tweedie, Markov chains and stochastic stability, 2nd Edition, Cambridge University Press, 2009.

[6] S. Louhichi, B. Ycart, Exponential growth of bifurcating processes with ancestral dependence, arXiv:1307.0343, submitted (2013).

[7] M. L. Kleptsyna, A. Le Breton, M. Viot, New formulas concerning Laplace transforms of quadratic forms for general Gaussian sequences, J. Appl. Math. Stoch. Anal. 15 (4) (2002) 309–325.